

Decoupling multivariate polynomials: interconnections between tensorizations

Konstantin Usevich^{a,b,*}, Philippe Dreesen^c, Mariya Ishteva^c

^a*Univ. Grenoble Alpes, GIPSA-Lab, F-38000 Grenoble, France*

^b*CNRS, GIPSA-Lab, F-38000 Grenoble, France*

^c*Vrije Universiteit Brussel (VUB), Department VUB-ELEC, Brussels, Belgium*

Abstract

Decoupling multivariate polynomials is a useful tool for obtaining an insight into the workings of a nonlinear mapping, to perform parameter reduction, or to approximate nonlinear functions. Several different tensor-based approaches have been proposed for this task, involving different tensor representations of the functions, and ultimately lead to a canonical polyadic decomposition. We will describe and study the connections between the different tensorization approaches, and connect two specific tensorization methods, such that an exchange of theoretical results between the two approaches is made possible. The first approach constructs a tensor that contains the coefficients of the polynomials, whereas the second approach builds a tensor from the Jacobian matrices of the polynomial vector function, evaluated at a set of sampling points. Although both approaches clearly share certain elements, they are quite different. In the current article, we study the connections and suggest ways to exploit the underlying structure of the employed tensorizations. We have found that the two approaches can be related to one another: the tensor containing the Jacobian matrices is obtained by multiplying the coefficient-based tensor with a Vandermonde-like matrix. The particular choice of tensorization approach has practical repercussions, and favors or disfavors easier computations or the implementation of certain aspects. The discussed connections provide a way to better assess which of the methods should be favored in certain problem settings, and may be a starting point to unify the two approaches.

Keywords: polynomial decoupling, tensors, quasi-Hankel matrices, system identification, Waring decomposition, tensorization, canonical polyadic decomposition, structured data fusion

2010 MSC: 12E05; 15A21; 15A69

1. Introduction

Representing a nonlinear function in a simpler way can be useful for providing an insight into its inner workings, to reduce the parametric complexity, or to perform func-

*Corresponding author

Email addresses: konstantin.usevich@gipsa-lab.fr (Konstantin Usevich),
philippe.dreesen@vub.ac.be (Philippe Dreesen), mariya.ishteva@vub.ac.be (Mariya Ishteva)

tion approximation. For this reason, the task of decoupling a set of polynomial vector functions, that is, to decompose a set of multivariate real polynomials into linear combinations of univariate polynomials in linear forms of the input variables, has attracted a spark of research attention over the last years. Restricting this question to homogeneous polynomials leads to the well-known Waring decomposition [1, 2, 3], but generalizations to non-homogeneous polynomials or the joint Waring decomposition are studied as well [4, 5] and [6, 7, 8]. The same type of decomposition appears in other fields, such as approximation theory [9, 10, 11], and polynomial neural networks [12]. It has also been applied successfully in the field of system identification [13, 14, 8, 15, 16, 17].

The polynomial decoupling problem has led to several tensor-based solution methods [15, 18, 19, 7]. These approaches can be categorized into two classes. The methods [15, 16, 18, 19] build a tensor from the polynomial coefficients, whereas the method of [7] builds a tensor from the Jacobian matrices of the functions, evaluated at a set of sampling points. All methods perform (in some way) a canonical polyadic decomposition (CP decomposition) of the constructed tensor to retrieve a decoupled representation in which the nonlinearities occur as univariate polynomial mappings.

The advantages of using a tensor-based approach for decoupling are in two aspects. First, by solving the decoupling problem as a CP decomposition, one can use recent widely available and robust numerical tools, such as Tensorlab for MATLAB [20] (or alternatives [21, 22]). Second, in general, ‘tensorization’ procedures are often able to take advantage of the notions of rank and the beneficial uniqueness properties of tensor representations or their CP decomposition [23], e.g., ensuring that a decoupling question is guaranteed to be identifiable.

In this article, we will specifically focus on the two tensorization methods [19] and [7], and explore their connections. Both approaches seem quite different in nature, although both associated tensors have a particular structure, and each of the methods has distinct advantages over the other one. For instance, the coefficient-based methods [15, 16, 18, 19] require several high-order tensors (or their matricizations) for polynomials of high degrees, whereas [7] always uses a third-order tensor. Coefficient-based approaches can easily deal with single polynomials, whereas [7] would in that case not be able to take advantage of the uniqueness properties of the CP decomposition, as the *tensor* of Jacobian matrices is then a matrix composed of gradient vectors. On the other hand, the approach of [7] can be applied to non-polynomial functions, which may in some cases be of interest, e.g., in [13] a neural network was decoupled. Obtaining a deeper understanding of the connections between the solution approaches is of interest in a non-algebraic or non-polynomial context. We will explore how the different decoupling approaches can be related to one another and we discuss their properties.

The remainder of this article is organized as follows: Section 2 formalizes the problem statement. Section 3 discusses the construction of the tensors. Section 4 presents the relations between the two alternatives and associated computational aspects. Section 5 recalls links between polynomials and tensors. Section 6 presents the proofs of the main results. Section 7 draws the conclusions and points out open problems for future work.

Notation

Scalars are denoted by lowercase or uppercase letters. Vectors are denoted by lowercase boldface letters, e.g., **u**. Elements of a vector are denoted by lowercase letters

with an index as subscript, e.g., $\mathbf{x} = [x_1 \ \cdots \ x_m]^\top$. Matrices are denoted by uppercase boldface letters, e.g., \mathbf{V} . The entry in the i -th row and j -th column of a matrix \mathbf{V} is denoted by v_{ij} , and the matrix $\mathbf{V} \in \mathbb{R}^{m \times r}$ may be represented by its columns $\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_r]$. Tensors of order d are denoted by uppercase caligraphical letters, e.g., $\mathcal{T} \in \mathbb{R}^{n \times m \times N}$. The outer product is denoted by “ \circ ” and is defined as follows: For $\mathcal{T} = \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}$, the entry in position (i, j, k) is equal to $u_i v_j w_k$. The canonical polyadic (CP) decomposition expresses a tensor \mathcal{T} as a (minimal) sum of rank-one tensor terms [24, 25, 26] as $\mathcal{T} = \sum_{i=1}^R \mathbf{u}_i \circ \mathbf{v}_i \circ \mathbf{w}_i$, see Figure 1, and is sometimes denoted in a shorthand notation as $\mathcal{T} = \llbracket \mathbf{U}, \mathbf{V}, \mathbf{W} \rrbracket$. The CP rank r is defined as the (minimal) number of terms that is required to represent \mathcal{T} as a sum of r rank-one terms. To refer to elements of matrices or tensors, or subsets thereof, we may use MATLAB-like index notation (including MATLAB’s colon wildcard): for instance, the element at position (i, j, k, ℓ) of a fourth-order tensor \mathcal{T} can be represented by $\mathcal{T}_{i,j,k,\ell}$, and the second frontal slice of a third-order tensor \mathcal{T} is as the matrix $\mathcal{T}_{:, :, 2}$. The mode- n product is denoted by “ \bullet_n ” and is defined as follows. Let \mathcal{X} be a $I_1 \times I_2 \times \cdots \times I_N$ tensor, and let \mathbf{u} be a vector of length I_n , then we have $(\mathcal{X} \bullet_n \mathbf{u}^\top)_{i_1 \cdots i_{n-1} i_{n+1} \cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \cdots i_N} u_{i_n}$. Notice that the result is a tensor of order $N - 1$, as mode n is summed out. The Kronecker product is denoted by “ \otimes ”.

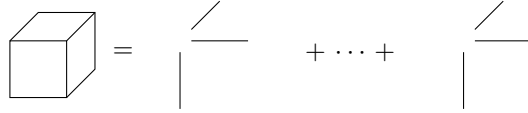


Figure 1: The canonical polyadic decomposition decomposes a third-order tensor into a minimal sum of rank-one terms, where each rank-one term is the outer product of three vectors. The number of terms is called the rank of the tensor.

2. A polynomial decoupling model

2.1. Problem statement

First, we describe the model, following the notation of [7] as illustrated in Figure 2. Consider a multivariate polynomial map $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$, i.e., a vector $\mathbf{f}(\mathbf{u}) = [f_1(\mathbf{u}) \ \cdots \ f_n(\mathbf{u})]^\top$ of multivariate polynomials (of total degree at most d) in variables $\mathbf{u} = [u_1 \ \cdots \ u_m]^\top$. We say that \mathbf{f} has a *decoupled representation*, if it can be expressed as

$$\mathbf{f}(\mathbf{u}) = \mathbf{W} \mathbf{g}(\mathbf{V}^\top \mathbf{u}), \quad (1)$$

where $\mathbf{V} \in \mathbb{R}^{m \times r}$ and $\mathbf{W} \in \mathbb{R}^{n \times r}$ are transformation matrices, and $\mathbf{g} : \mathbb{R}^r \rightarrow \mathbb{R}^n$ is defined as

$$\mathbf{g}(x_1, \dots, x_r) = [g_1(x_1) \ \cdots \ g_r(x_r)]^\top,$$

where $g_k : \mathbb{R} \rightarrow \mathbb{R}$ are univariate polynomials of degree at most d , i.e., $g_k(t) = c_1 t + \cdots + c_d t^d$. Note that we omitted the constant terms of the polynomials, since they are not uniquely identifiable [7]. In this paper we limit ourselves to the model (1), although other decoupling models can be considered.

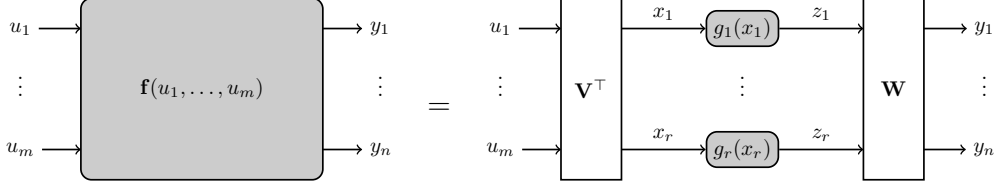


Figure 2: Every multivariate polynomial vector function $\mathbf{f}(\mathbf{u})$ can be represented by a linear transformation of a set of univariate functions (in linear combinations of the original variables).

The decoupled representation (1) can be also represented as an additive decomposition

$$\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u}), \quad (2)$$

where \mathbf{v}_k and \mathbf{w}_k are the columns of \mathbf{V} and \mathbf{W} respectively. As shown in [27, 28], the decomposition (2) is a special case of the X -rank decomposition, where the set of “rank-one” terms is the set of polynomial maps of the form $\mathbf{w}g(\mathbf{v}^\top \mathbf{u})$. The X -rank framework [28] is useful for studying the identifiability (i.e., generic uniqueness) of the model (2).

The following example shows a decoupled representation for a simple case. This example will be used throughout the paper to illustrate the main ideas of the various aspects that we will explore.

Example 1. Consider a function $\mathbf{f}(\mathbf{u}) = [f_1(u_1, u_2) \ f_2(u_1, u_2)]^\top$ given as

$$\begin{aligned} f_1(u_1, u_2) &= -3u_1^3 - 9u_1^2 u_2 - 27u_1 u_2^2 - 15u_2^3 - 8u_1^2 - 8u_1 u_2 - 20u_2^2 + 3u_1 + 9u_2, \\ f_2(u_1, u_2) &= -7u_1^3 - 6u_1^2 u_2 + 6u_1 u_2^2 + 7u_2^3 + 10u_1^2 + 16u_1 u_2 + 10u_2^2 - 3u_2. \end{aligned}$$

It can be verified that \mathbf{f} has a decomposition $\mathbf{f}(\mathbf{u}) = \mathbf{w}_1 g_1(\mathbf{v}_1^\top \mathbf{u}) + \cdots + \mathbf{w}_r g_r(\mathbf{v}_r^\top \mathbf{u})$ with $m = n = 2$ and $r = 3$ as

$$\mathbf{V} = \begin{bmatrix} 2 & -1 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \text{and} \quad \mathbf{W} = \begin{bmatrix} 0 & 1 & -2 \\ -1 & 0 & 1 \end{bmatrix},$$

and

$$\begin{aligned} g_1(x_1) &= x_1^3 - 2x_1^2 - x_1, \\ g_2(x_2) &= x_2^3 - 4x_2^2 + x_2, \\ g_3(x_3) &= x_3^3 + 2x_3^2 - 2x_3. \end{aligned}$$

3. Tensorizations and their decompositions

In this section, we recall the tensorizations of [19] and [7] and discuss their basic properties. For completeness, we give short proofs of the specific structures of the corresponding CP decompositions, although these proofs are already present in [19, 7].

3.1. Simultaneous coupled symmetric CP decomposition

Let us review some facts that connect polynomials with symmetric tensor representation [29, 30, 31]. This will lead naturally to the tensorization method of [19], which will play an important role in the remainder of the paper.

Polynomials and symmetric tensors

It is well-known that a homogeneous polynomial can be represented as a symmetric tensor [31]. For instance, the polynomial $-8u_1^2 + -8u_1u_2 - 20u_2^2$ can be written using a symmetric coefficient matrix $\Psi^{(2)} \in \mathbb{R}^{2 \times 2}$ as

$$-8u_1^2 + -8u_1u_2 - 20u_2^2 = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} -8 & -4 \\ -4 & -20 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \mathbf{u}^\top \Psi^{(2)} \mathbf{u}. \quad (3)$$

Any non-homogeneous polynomial of degree d can hence be written as

$$p(\mathbf{u}) = \mathbf{u}^\top \Psi^{(1)} + \mathbf{u}^\top \Psi^{(2)} \mathbf{u} + \Psi^{(3)} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \bullet_3 \mathbf{u} + \dots + \Psi^{(d)} \bullet_1 \mathbf{u} \dots \bullet_d \mathbf{u}, \quad (4)$$

where $\Psi^{(1)} \in \mathbb{R}^m$, $\Psi^{(2)} \in \mathbb{R}^{m \times m}$ is a symmetric matrix, and $\Psi^{(d)} \in \mathbb{R}^{m \times \dots \times m}$ are symmetric tensors of order d .

Example 2. We continue Example 1. We can write $f_1(u_1, u_2)$ and $f_2(u_1, u_2)$ as

$$f_1(u_1, u_2) = \mathbf{u}^\top \begin{bmatrix} 3 \\ 9 \end{bmatrix} + \mathbf{u}^\top \begin{bmatrix} -8 & -4 \\ -4 & -20 \end{bmatrix} \mathbf{u} + \Psi^{(3)} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \bullet_3 \mathbf{u},$$

with

$$\Psi_{::,1}^{(3)} = \begin{bmatrix} -3 & -3 \\ -3 & -9 \end{bmatrix}, \quad \Psi_{::,2}^{(3)} = \begin{bmatrix} -3 & -9 \\ -9 & -15 \end{bmatrix},$$

and

$$f_2(u_1, u_2) = \mathbf{u}^\top \begin{bmatrix} 0 \\ -3 \end{bmatrix} + \mathbf{u}^\top \begin{bmatrix} 10 & 8 \\ 8 & 10 \end{bmatrix} \mathbf{u} + \Psi^{(3)} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \bullet_3 \mathbf{u},$$

with

$$\Psi_{::,1}^{(3)} = \begin{bmatrix} -7 & -2 \\ -2 & 2 \end{bmatrix}, \quad \Psi_{::,2}^{(3)} = \begin{bmatrix} -2 & 2 \\ 2 & 7 \end{bmatrix}.$$

Decoupling a single homogeneous polynomial

Before we look at the general case of decoupling several non-homogeneous polynomials, we first consider a single homogeneous polynomial. Let $p(u_1, \dots, u_m)$ denote a d -th degree polynomial in m variables, which can be represented by a symmetric d -th order tensor $\Psi^{(d)}$. The symmetric CP decomposition of $\Psi^{(d)}$ is closely related to its Waring decomposition [1, 2, 3, 32], describing how the polynomial $p(u_1, \dots, u_m)$ is written as a sum of d -th degree powers of linear forms of the input variables. Following our notation of the decoupling problem (1), the Waring decomposition for a degree d homogeneous polynomial takes the form

$$p(u_1, \dots, u_m) = \sum_{i=1}^r w_i (v_{1i}u_1 + \dots + v_{mi}u_m)^d. \quad (5)$$

The symmetric CP decomposition of $\Psi^{(d)}$ reveals possible values for the unknowns v_{ij} and w_i [29].

Example 3. Consider the second degree part of $f_1(u_1, u_2)$ from Example 2. The symmetric coefficient matrix $\Psi^{(2)}$ admits the decomposition

$$\begin{bmatrix} -8 & -4 \\ -4 & -20 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & 4 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -2 & 2 \\ 2 & 4 \end{bmatrix}, \quad (6)$$

such that $p(u_1, u_2) = \mathbf{u}^\top \Psi^{(2)} \mathbf{u}$ has the Waring decomposition

$$p(u_1, u_2) = -(-2u_1 + 2u_2)^2 - (2u_1 + 4u_2)^2.$$

Notice that the symmetric decomposition of $\Psi^{(2)}$ is not unique (nor ‘essentially unique’ [26]): its eigenvalue decomposition provides another valid factorization.

Decoupling several homogeneous polynomials

Along the same lines, it is possible to perform a joint decoupling of several homogeneous polynomials. Consider the case of n homogeneous polynomials of degree d , denoted by

$$\begin{aligned} p_1(u_1, \dots, u_m) &= \Psi_1^{(d)} \bullet_1 \mathbf{u} \cdots \bullet_d \mathbf{u}, \\ &\vdots \\ p_n(u_1, \dots, u_m) &= \Psi_n^{(d)} \bullet_1 \mathbf{u} \cdots \bullet_d \mathbf{u}. \end{aligned} \quad (7)$$

Let us arrange the $\Psi_i^{(d)}$, for $i = 1, \dots, n$ into a tensor \mathcal{T}^d , such that $\mathcal{T}_{i, \dots, i}^d = \Psi_i^{(d)}$, for all $i = 1, \dots, n$. Then it is easy to verify that this tensor \mathcal{T}^d admits a partially symmetric CP decomposition as

$$\mathcal{T}^d = \llbracket \mathbf{W}, \underbrace{\mathbf{V}, \dots, \mathbf{V}}_{d \text{ times}} \rrbracket,$$

which, in our decoupling representation (1), takes the form $\mathbf{W}\mathbf{g}(\mathbf{V}^\top \mathbf{u})$, where $\mathbf{g}(\mathbf{x}) = [x_1^d \cdots x_r^d]^\top$, and \mathbf{W} serves to make a linear combination of all the r internal branches for each of the n function components.

Decoupling several non-homogeneous polynomials

Finally, several non-homogeneous polynomials can be jointly decomposed in a similar way. Consider n non-homogeneous polynomials of maximal degree d , denoted as

$$\begin{aligned} p_1(u_1, \dots, u_m) &= \mathbf{u}^\top \Psi_1^{(1)} + \cdots + \Psi_1^{(d)} \bullet_1 \mathbf{u} \cdots \bullet_d \mathbf{u}, \\ &\vdots \\ p_n(u_1, \dots, u_m) &= \mathbf{u}^\top \Psi_n^{(1)} + \cdots + \Psi_n^{(d)} \bullet_1 \mathbf{u} \cdots \bullet_d \mathbf{u}, \end{aligned} \quad (8)$$

Let us arrange all $\Psi_i^{(d)}$, for $i = 1, \dots, n$ into the tensors \mathcal{T}^d (like in the previous paragraph), such that $\mathcal{T}_{i, \dots, i}^d = \Psi_i^{(d)}$, for all $i = 1, \dots, n$. We now have for each degree a coupled partially symmetric CP decomposition as

$$\begin{aligned} \mathcal{T}^1 &= \llbracket \mathbf{W}, \mathbf{V}, \mathbf{c}_1^\top \rrbracket, \\ \mathcal{T}^2 &= \llbracket \mathbf{W}, \mathbf{V}, \mathbf{V}, \mathbf{c}_2^\top \rrbracket, \\ &\vdots \\ \mathcal{T}^d &= \llbracket \mathbf{W}, \underbrace{\mathbf{V}, \dots, \mathbf{V}}_{d \text{ times}}, \mathbf{c}_d^\top \rrbracket, \end{aligned} \quad (9)$$

where the \mathbf{c}_i , for $i = 1, \dots, d$, are the i -th degree coefficients for each of the r branches.

It is worth mentioning that the joint decomposition of the coefficient tensors of several degrees is very similar to the ideas developed in [18, 16].

Remark that these coefficients were not required in the previous paragraphs when homogeneous polynomials were considered: in such cases only the scaling of the r branches contains a single scalar, which can be assumed to be fully absorbed by \mathbf{W} . Also remark that there is a redundancy in the representation (9): An equivalent problem is obtained if one of the coefficient vectors \mathbf{c}_δ is omitted (i.e., assumed to be equal to a vector containing ones), in which case a rescaling has to take place on the remaining coefficients as well as on \mathbf{W} . Finally we want to mention that the framework of structured data fusion [33, 20] allows for computing tensor decompositions as in (9), where several tensors (and possibly matrices) are jointly decomposed while sharing factors, possibly while imposing structure on the factors [20].

3.2. Tensor of unfoldings [19]

The above link between polynomials, (partially) symmetric tensors and their CP decompositions gives rise to the tensorization approach of [19], in which a tensor is constructed from the coefficients of the polynomials $f_1(u_1, \dots, u_m)$ up to $f_n(u_1, \dots, u_m)$. This tensorization offers the advantage that several polynomials can be represented as a single tensor, and the decoupling task can be solved using a single (but structured) CP decomposition. In this approach, the tensor (shown in Figure 3) is constructed from the coefficients of the polynomial map of degree d , as follows:

- The tensor has size $n \times m \times \delta$, where $\delta = \sum_{k=1}^d m^{k-1}$.
- The tensor is constructed by slices

$$\mathcal{Q}_{i,:,:} := \Psi(f_i),$$

where Ψ is a structured $m \times \delta$ matrix built from the coefficients of $f_i(\mathbf{u})$.

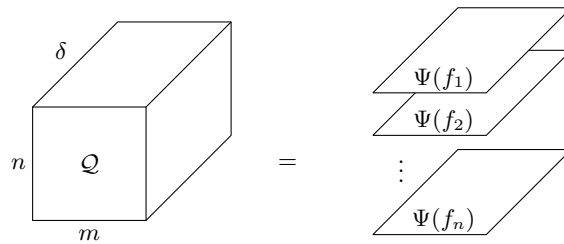


Figure 3: The coefficients of a polynomial map $\mathbf{f} : \mathbb{R}^m \rightarrow \mathbb{R}^n$ of degree d can be arranged into an $n \times m \times \delta$ tensor \mathcal{Q} , where $\delta = \sum_{k=1}^d m^{k-1}$.

Now let us describe the construction of the structured coefficient matrix $\Psi(p)$ for a given polynomial of degree d . Recall from (4) that each such polynomial can be written

as in (4), where $\Psi^{(1)} \in \mathbb{R}^m$, $\Psi^{(2)} \in \mathbb{R}^{m \times m}$ is a symmetric matrix and $\Psi^{(s)} \in \mathbb{R}^{m \times \dots \times m}$ are symmetric tensors of order s . Then the matrix $\Psi(p) \in \mathbb{R}^{m \times \delta}$ is constructed¹ as

$$\Psi(p) = \left[\begin{array}{c|c|c|c|c} \Psi^{(1)} & \Psi^{(2)} & \Psi^{(3)} & \dots & \Psi^{(d)} \\ \hline \end{array} \right],$$

where $\mathcal{G}_{(1)}$ denotes the first-mode unfolding of a tensor \mathcal{G} .

Example 4. A third-degree polynomial in two variables

$$p(u_1, u_2) = a_1 u_1 + a_2 u_2 + b_1 u_1^2 + 2b_2 u_1 u_2 + b_3 u_2^2 + d_1 u_1^3 + 3d_2 u_1^2 u_2 + 3d_3 u_1 u_2^2 + d_4 u_2^3$$

has the representation

$$p(u_1, u_2) = \mathbf{u}^\top \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \mathbf{u}^\top \begin{bmatrix} b_1 & b_2 \\ b_2 & b_3 \end{bmatrix} \mathbf{u} + \Psi^{(3)} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \bullet_3 \mathbf{u}, \quad (10)$$

where

$$\Psi_{::,1}^{(3)} = \begin{bmatrix} d_1 & d_2 \\ d_2 & d_3 \end{bmatrix}, \quad \Psi_{::,2}^{(3)} = \begin{bmatrix} d_2 & d_3 \\ d_3 & d_4 \end{bmatrix}.$$

By putting all the unfoldings together, we get

$$\Psi(p) = \left[\begin{array}{c|c|c|c|c} a_1 & b_1 & b_2 & d_1 & d_2 & d_2 & d_3 & d_3 \\ \hline a_2 & b_2 & b_3 & d_2 & d_3 & d_3 & d_3 & d_4 \end{array} \right]. \quad (11)$$

Hence, for the polynomials f_1 and f_2 in Example 1, the slices of the tensor \mathcal{Q} are given by

$$\mathcal{Q}_{1,::} = \Psi(f_1) = \left[\begin{array}{c|c|c|c|c} 3 & -8 & -4 & -3 & -3 & -3 & -9 \\ \hline 9 & -4 & -20 & -3 & -9 & -9 & -15 \end{array} \right],$$

and

$$\mathcal{Q}_{2,::} = \Psi(f_2) = \left[\begin{array}{c|c|c|c|c} 0 & 10 & 8 & -7 & -2 & -2 & 2 \\ \hline -3 & 8 & 10 & -2 & 2 & 2 & 7 \end{array} \right].$$

It appears that the tensor \mathcal{Q} has a CP decomposition, which reveals the decomposition (1), as proved in [19]. We repeat here a simplified version of the proof for completeness.

Lemma 1. For the polynomial map (1), the tensor \mathcal{Q} has the following CP decomposition:

$$\mathcal{Q} = \sum_{k=1}^r \mathbf{w}_k \circ \mathbf{v}_k \circ \mathbf{z}_k, \quad (12)$$

where

$$\mathbf{z}_k = \left[\begin{array}{c|c|c|c|c} c_{k,1} & c_{k,2} \mathbf{v}_k^\top & c_{k,3} (\mathbf{v}_k \otimes \mathbf{v}_k)^\top & \dots & c_{k,d} (\mathbf{v}_k \otimes \dots \otimes \mathbf{v}_k)^\top \end{array} \right]^\top. \quad (13)$$

¹In the original paper [19] the linear term is skipped, and $\delta = \sum_{k=2}^d m^{d-1}$. In [34, Appendix A.2] the matrix is denoted as Γ .

Proof. Consider the polynomial $q_k(\mathbf{u}) := g_k(\mathbf{v}_k^\top \mathbf{u})$, where $g_k(t) = c_{k,1}t + \dots + c_{k,d}t^d$, easy calculations show that

$$\Psi(q_k) = \mathbf{v}_k \mathbf{z}_k^\top,$$

see also [34, eqn. (A.7)]. Since, from (2) $f_i(\mathbf{u}) = \sum_{k=1}^r (\mathbf{w}_k)_i q_k(\mathbf{u})$, we have that

$$\Psi(f_i) = \sum_{k=1}^r (\mathbf{w}_k)_i \mathbf{v}_k \mathbf{z}_k^\top$$

which implies (12). \square

Remark that in Lemma 1 we see that the structure appearing in the CP decomposition of \mathcal{Q} is closely connected to the simultaneous decomposition described in Section 3.1. Indeed, the tensor \mathcal{Q} is equivalent to reshaping the tensors in (9) into third-order tensors, and stacking them along the third mode together.

Example 5. We continue Examples 1 and 4. The Kronecker products of the columns of \mathbf{V} are:

$$(\mathbf{v}_1 \otimes \mathbf{v}_1)^\top = [4 \ 2 \ 2 \ 1], (\mathbf{v}_2 \otimes \mathbf{v}_2)^\top = [1 \ -1 \ -1 \ 1], (\mathbf{v}_3 \otimes \mathbf{v}_3)^\top = [1 \ 2 \ 2 \ 4].$$

Hence, the matrix $\mathbf{Z} = [\mathbf{z}_1 \ \mathbf{z}_2 \ \mathbf{z}_3]$ is given by

$$\mathbf{Z}^\top = \left[\begin{array}{c|c|c} -1 & -4 & -2 & 4 & 2 & 2 & 1 \\ 1 & 4 & -4 & 1 & -1 & -1 & 1 \\ -2 & 2 & 4 & 1 & 2 & 2 & 4 \end{array} \right].$$

3.3. The tensor of Jacobian matrices of [7]

The tensorization method of [7] does not use the coefficients of $\mathbf{f}(\mathbf{u})$ directly, but is inspired on a small-signal analysis of the nonlinear function about a set of operating points. The method proceeds by collecting the first-order information of $\mathbf{f}(\mathbf{u})$ (i.e., the partial derivatives) in a set of sampling points. The thusly obtained Jacobian matrices are arranged into a third-order tensor, of which the CP decomposition reveals the decomposition (1).

As in [7], we consider the Jacobian of \mathbf{f} :

$$\mathbf{J}_{\mathbf{f}}(\mathbf{u}) := \begin{bmatrix} \frac{\partial f_1}{\partial u_1}(\mathbf{u}) & \dots & \frac{\partial f_1}{\partial u_m}(\mathbf{u}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial u_1}(\mathbf{u}) & \dots & \frac{\partial f_n}{\partial u_m}(\mathbf{u}) \end{bmatrix}. \quad (14)$$

Using Lemma 2, the tensorization is constructed as follows (see Figure 4):

- N points $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(N)} \in \mathbb{R}^m$ are chosen (so-called *sampling points*).
- An $n \times m \times N$ tensor is constructed by stacking the Jacobian evaluations at the sampling points

$$\mathcal{J}_{::,k} := \mathbf{J}_{\mathbf{f}}(\mathbf{u}^{(k)}).$$

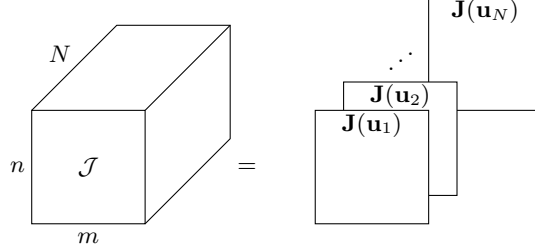


Figure 4: The third-order tensor \mathcal{J} is constructed by stacking behind each other a set of Jacobian matrices \mathbf{J} evaluated at the sampling points $\mathbf{u}^{(k)}$. Its CP decomposition is equivalent to joint matrix diagonalization of the Jacobian matrix slices.

Example 6. We continue Example 1. The partial derivatives of $\mathbf{f}(\mathbf{u})$ that form the Jacobian matrix are

$$\begin{aligned}\frac{\partial f_1}{\partial u_1} &= -9u_1^2 - 18u_1u_2 - 27u_2^2 - 16u_1 - 8u_2 + 3, \\ \frac{\partial f_1}{\partial u_2} &= -9u_1^2 - 54u_1u_2 - 45u_2^2 - 8u_1 - 40u_2 + 9, \\ \frac{\partial f_2}{\partial u_1} &= -21u_1^2 - 12u_1u_2 + 6u_2^2 + 20u_1 + 16u_2, \\ \frac{\partial f_2}{\partial u_2} &= -6u_1^2 + 12u_1u_2 + 21u_2^2 + 16u_1 + 20u_2 - 3.\end{aligned}$$

As a set of sampling points, we choose

$$\mathbf{u}^{(1)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \mathbf{u}^{(2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{u}^{(3)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

By evaluating $\mathbf{J}_{\mathbf{f}}(\mathbf{u})$ at these points, we get the tensor $\mathcal{J}^{(1)} = \mathcal{J}$ given by

$$\mathcal{J}_{::,1}^{(1)} = \begin{bmatrix} 3 & 9 \\ 0 & -3 \end{bmatrix}, \quad \mathcal{J}_{::,2}^{(1)} = \begin{bmatrix} -22 & -8 \\ -1 & 7 \end{bmatrix}, \quad \mathcal{J}_{::,3}^{(1)} = \begin{bmatrix} -32 & -76 \\ 22 & 38 \end{bmatrix}.$$

Next, we show that if $\mathbf{f}(\mathbf{u})$ has a decoupled representation (1), then the following lemma holds true.

Lemma 2 ([7, Lemma 2.1]). The first order derivatives of (1) are given by

$$\mathbf{J}_{\mathbf{f}}(\mathbf{u}) = \mathbf{W} \text{diag}(g'_1(\mathbf{v}_1^\top \mathbf{u}), \dots, g'_r(\mathbf{v}_r^\top \mathbf{u})) \mathbf{V}^\top, \quad (15)$$

where $g'_i(t) := \frac{dg_i}{dt}(t)$.

Sketch of the proof. By the chain rule for multivariate functions, we have

$$\mathbf{J}_{\mathbf{f}}(\mathbf{u}) = \mathbf{W} \mathbf{J}_{\mathbf{g}}(\mathbf{V}^\top \mathbf{u}) \mathbf{V}^\top,$$

which is exactly (15). □

By Lemma 2, the evaluations of the Jacobians can be jointly factorized:

$$\begin{aligned} \mathbf{J}(\mathbf{u}^{(1)}) &= \mathbf{W}\mathbf{D}^{(1)}\mathbf{V}^\top, \\ &\vdots \\ \mathbf{J}(\mathbf{u}^{(N)}) &= \mathbf{W}\mathbf{D}^{(N)}\mathbf{V}^\top, \end{aligned} \tag{16}$$

where $\mathbf{D}^{(k)} = \text{diag}(d_1^{(k)}, \dots, d_r^{(k)})$ and $d_i^{(k)} = g'_i(\mathbf{v}_i^\top \mathbf{u}^{(k)})$. Therefore, the tensor \mathcal{J} admits the following CP decomposition

$$\mathcal{J} = \sum_{k=1}^r \mathbf{w}_k \circ \mathbf{v}_k \circ \mathbf{h}_k, \tag{17}$$

where vectors $\mathbf{w}_k, \mathbf{v}_k$ are as defined earlier, and \mathbf{h}_k contains the evaluations of $g'_k(\mathbf{v}_k^\top \mathbf{u})$ in N sampling points

$$\mathbf{h}_k = [g'_k(\mathbf{v}_k^\top \mathbf{u}^{(1)}) \quad \dots \quad g'_k(\mathbf{v}_k^\top \mathbf{u}^{(N)})]. \tag{18}$$

Example 7. We continue Examples 1 and 6. By differentiation, we get

$$\begin{aligned} g'_1(t) &= 3t^2 - 4t - 1, \\ g'_2(t) &= 3t^2 - 8t + 1, \\ g'_3(t) &= 3t^2 + 4t - 2, \end{aligned}$$

and hence, by substitution,

$$\mathbf{H} = \begin{bmatrix} -1 & 1 & -2 \\ 3 & 12 & 5 \\ -2 & -4 & 18 \end{bmatrix}.$$

Straightforward calculations show indeed that $\mathcal{T}^{(1)}$ admits a decomposition (17).

4. Properties of tensorizations

In this section, we show how the tensor decompositions of (12) and (17) are related. Moreover, we establish the relation between the ranks of the decompositions and their uniqueness properties.

4.1. Relation between \mathcal{J} and \mathcal{Q}

First, we show the relation between the vectors defined in (13) and (18).

Lemma 3. The vectors \mathbf{h}_k and \mathbf{z}_k defined in (13) and (18), respectively, satisfy the following relation:

$$\mathbf{h}_k = \mathbf{A}^\top \mathbf{z}_k, \tag{19}$$

where $\mathbf{A} \in \mathbb{R}^{\delta \times N}$ is a Vandermonde-like matrix whose columns are

$$\mathbf{A}_{:,k} = [1 \mid 2(\mathbf{u}^{(k)})^\top \mid 3(\mathbf{u}^{(k)} \otimes \mathbf{u}^{(k)})^\top \mid \dots \mid d(\mathbf{u}^{(k)} \otimes \dots \otimes \mathbf{u}^{(k)})^\top]^\top. \tag{20}$$

Proof. Let us express $g'_k(\mathbf{v}_k^\top \mathbf{u})$ in an explicit form. First, $g'_k(t) = c_{k,1} + 2c_{k,2}t + 3c_{k,3}t^2 + \dots + dc_{k,d}t^{d-1}$, from which it follows that

$$g'_k(\mathbf{v}_k^\top \mathbf{u}) = c_{k,1} + 2c_{k,2}\mathbf{v}_k^\top \mathbf{u} + 3c_{k,3}(\mathbf{v}_k^\top \mathbf{u})^2 + \dots + dc_{k,d}(\mathbf{v}_k^\top \mathbf{u})^{d-1}.$$

Since $(\mathbf{v}^\top \mathbf{u})^s = (\mathbf{v} \otimes \dots \otimes \mathbf{v})^\top (\mathbf{u} \otimes \dots \otimes \mathbf{u})$, we have that the j -th element of \mathbf{h}_k is equal to

$$(\mathbf{h}_k)_j = h_{j,k} = g'_k(\mathbf{v}_k^\top \mathbf{u}^{(j)}) = \mathbf{A}_{:,k}^\top \mathbf{z}_k,$$

which completes the proof. \square

As a consequence, we get that the two tensors are also related.

Proposition 1. *For any polynomial map \mathbf{f} , the tensors \mathcal{J} and \mathcal{Q} are related as*

$$\mathcal{J} = \mathcal{Q} \bullet_3 \mathbf{A}^\top. \quad (21)$$

Proof. First, any polynomial map \mathbf{f} can be decomposed as (1) with r sufficiently large. Let us take such a decomposition; then it holds that

$$(\mathcal{Q}) \bullet_3 \mathbf{A}^\top = \left(\sum_{k=1}^r \mathbf{w}_k \circ \mathbf{v}_k \circ \mathbf{z}_k \right) \bullet_3 \mathbf{A}^\top = \sum_{k=1}^r \mathbf{w}_k \circ \mathbf{v}_k \circ \mathbf{A}^\top \mathbf{z}_k = \mathcal{J},$$

where the last inequality follows from (19). \square

Example 8. *In Example 6, the matrix \mathbf{A} can be found as*

$$\mathbf{A}^\top = \left[\begin{array}{c|cc|ccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 3 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 0 & 0 & 3 \end{array} \right].$$

It is easy to see that $\mathbf{H} = \mathbf{A}^\top \mathbf{Z}$.

4.2. Compressed version of \mathcal{Q}

Although Proposition 1 establishes the connection between the tensors \mathcal{J} and \mathcal{Q} , it does not say anything about connections of their ranks and minimal decompositions. One of the obvious obstacles is that both \mathcal{Q} and the matrix \mathbf{A} defined in (20) have repeating rows. We are going to remedy this problem in the following proposition by introducing a compressed representation of these matrices and tensors.

Proposition 2. *Let $M = \binom{m+d-1}{d-1}$. There exists a selection matrix $\mathbf{S} \in \{0,1\}^{\delta \times M}$, that is, a $\delta \times M$ matrix containing only zeros and ones, such that each row contains one single non-zero element, and \mathbf{S} has full column rank. Moreover, there exists a compressed tensor $\mathcal{C} \in \mathbb{R}^{n \times m \times M}$ and matrix $\mathbf{B} \in \mathbb{R}^{M \times r}$ such that*

$$\mathcal{Q} = \mathcal{C} \bullet_3 \mathbf{S}, \quad \mathbf{A} = \mathbf{S}\mathbf{B}.$$

We leave the proof of the proposition for Section 6, but already can give an example of such a selection matrix.

Example 9. In Example 4, the selection matrix is

$$\mathbf{S} = \begin{bmatrix} \mathbf{I}_4 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}$$

Finally, the selection matrix helps us to establish some statements about ranks of the tensors

Proposition 3.

1. If the matrix \mathbf{B} has full row rank, then the CP ranks of \mathcal{J} , \mathcal{Q} and \mathcal{C} coincide. Moreover, if one of the CP decompositions is unique, then all the others are also unique.
2. If $N \geq M$, and the sampling points are in general position, then the matrix \mathbf{B} has full row rank.

The proof will be given in Section 6.

4.3. Structured tensor rank

Although, as shown in the previous subsection, it is possible to remove repeating rows and columns of \mathcal{Q} , there are still nontrivial linear dependencies between the elements of \mathcal{Q} . Hence, the CP decomposition of \mathcal{J} and \mathcal{Q} is, in general, not equivalent to the original decomposition (2). In what follows, we establish relationships between the CP decompositions and the original decomposition (2).

First, we prove that for the rank-one case, these decompositions coincide.

Proposition 4. Consider a polynomial map $\mathbf{f}(\mathbf{u})$ of degree d , and the tensor \mathcal{Q} built from it. Then the following holds

$$\text{rank}(\mathcal{Q}) = \text{rank}(\mathcal{C}) \leq 1 \iff \mathbf{f}(\mathbf{u}) = \mathbf{w}g(\mathbf{v}^\top \mathbf{u}),$$

where $\mathbf{w} \in \mathbb{R}^n$, $\mathbf{v} \in \mathbb{R}^m$ and $g(t)$ is a polynomial of degree d .

The proof is given in Section 6.

Corollary 1. If the N sampling points are chosen such that the matrix $\text{rank}(\mathbf{A}) = M$, then

$$\text{rank}(\mathcal{J}) \leq 1 \iff \mathbf{f}(\mathbf{u}) = \mathbf{w}g(\mathbf{v}^\top \mathbf{u}).$$

As a corollary of Proposition 4, we get that the original polynomial decomposition (2) is equivalent to a structured CP decomposition.

Corollary 2. Let $\mathcal{L}_{\mathcal{Q}} \subset \mathbb{R}^{n \times m \times \delta}$ be the linear subspace of tensors with the structure of \mathcal{Q} . Let the sampling points be chosen such that $\text{rank} \mathbf{A} = M$, and $\mathcal{L}_{\mathcal{J}} \subset \mathbb{R}^{n \times m \times N}$ be the linear subspace of tensors with the structure of \mathcal{J} .

Then the following three statements are equivalent:

1. the polynomial map $\mathbf{f}(\mathbf{u})$ admits a decomposition (2);

2. the tensor $\mathcal{Q}(\mathbf{f})$ admits the structured CP decomposition

$$\mathcal{Q} = \mathcal{Q}_1 + \cdots + \mathcal{Q}_r, \quad \text{rank}(\mathcal{Q}_k) = 1, \quad \mathcal{Q}_k \in \mathcal{L}_{\mathcal{Q}}; \quad (22)$$

3. the tensor $\mathcal{J}(\mathbf{f})$ admits the structured CP decomposition

$$\mathcal{J} = \mathcal{J}_1 + \cdots + \mathcal{J}_r, \quad \text{rank}(\mathcal{J}_k) = 1, \quad \mathcal{J}_k \in \mathcal{L}_{\mathcal{J}}. \quad (23)$$

The structure constraint is important: indeed, the CP decomposition of the tensor \mathcal{Q} or \mathcal{J} is not necessarily structured. In general, we do not know even if the CP rank is equal to the structured CP rank (minimal number of terms in (22) or (23)). This is similar to the Comon's conjecture about symmetric tensors: it is not known whether the symmetric rank of a symmetric tensor equals its non-symmetric rank.

However, if the CP decomposition of a tensor is unique (for example, if it satisfies Kruskal's uniqueness conditions), then it should necessarily be a structured CP decomposition.

Example 10. Let us continue with Examples 1, 6 and 7. We have that $m = n = 2$ and $r = 3$, which does not guarantee a unique CP decomposition of \mathcal{J} (under assumptions of genericity, see [7]). Indeed, if we compute a numerical CP decomposition of tensor \mathcal{J} , we find that, up to a relative norm-wise error 2.3546×10^{-16} , \mathcal{J} admits a decomposition with factors

$$\begin{aligned} \tilde{\mathbf{W}} &= \begin{bmatrix} 1.1628 & -3.2951 & 3.0252 \\ 0.5705 & 1.1349 & -2.1791 \end{bmatrix}, \\ \tilde{\mathbf{V}} &= \begin{bmatrix} 3.5822 & -0.7705 & -2.2959 \\ -0.0226 & -3.4455 & -2.9785 \end{bmatrix}, \\ \tilde{\mathbf{H}} &= \begin{bmatrix} 0.2736 & 0.8181 & 0.0312 \\ -3.3900 & 0.2647 & 1.2313 \\ 0.5334 & -3.9194 & 3.4945 \end{bmatrix}, \end{aligned}$$

the columns of which are not scaled and permuted versions of the columns of \mathbf{W} , \mathbf{V} and \mathbf{H} .

It can be shown that the 'structured CP' approaches are able to correctly return the underlying factors \mathbf{W} , \mathbf{V} and \mathbf{H} (up to scaling and permutation invariances). For instance, the structured data fusion framework [33, 20] is able to compute the coupled and partially symmetric decomposition as it was presented in Section 3.1 in (9). This returns

$$\begin{aligned} \tilde{\mathbf{W}} &= \begin{bmatrix} 1.2767 & 1.7112 & 0 \\ 0 & -0.8556 & -1.9980 \end{bmatrix}, \\ \tilde{\mathbf{V}} &= \begin{bmatrix} -0.9218 & -1.0534 & 1.5879 \\ 0.9218 & -2.1067 & 0.7940 \end{bmatrix}, \end{aligned}$$

as well as computed values for the coefficient vectors of $g_i(x_i)$, which are omitted here. It can be verified that the returned factors \mathbf{W} and \mathbf{V} are scaled and permuted versions of the factors \mathbf{W} and \mathbf{V} .

Remark that if one uses $m = n = r = 2$, both methods return the same CP decomposition (up to scaling and permutation of the columns of the factors). Indeed, in this case,

uniqueness is guaranteed (generically), ensuring that the underlying factors are identifiable. This could be checked easily by generating a variation of the equations that we are decoupling where the third columns of \mathbf{V} and \mathbf{W} are removed, so that $g_3(x_3)$ is not considered.

5. Tensors, polynomials and quasi-Hankel matrices

In this section, we recall some preliminaries on polynomials and tensors, which will be later used in Section 6.

5.1. Polynomials and multinomial coefficients

For a multi-index $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{Z}_+^m$, we use the following short-hand notation for monomials:

$$\mathbf{u}^\alpha = u_1^{\alpha_1} \dots u_m^{\alpha_m},$$

and

$$|\alpha| = \alpha_1 + \dots + \alpha_m,$$

which is the total degree of \mathbf{u}^α . Next, for a multidegree $\alpha = (\alpha_1, \dots, \alpha_m)$ we define the multinomial coefficient as

$$\binom{|\alpha|}{\alpha} = \frac{(\alpha_1 + \dots + \alpha_m)!}{\alpha_1! \dots \alpha_m!}.$$

Finally, assume that a polynomial of degree d in m variables $f(\mathbf{u})$ is given in a basis of monomials scaled by multinomial coefficients.

$$f(\mathbf{u}) = \sum_{\alpha_1, \dots, \alpha_m=0}^{|\alpha| \leq d} \binom{|\alpha|}{\alpha} f_\alpha \mathbf{u}^\alpha. \quad (24)$$

Example 11. Let $m = 2$. In this case, the multinomial coefficients are binomial coefficients, i.e. $\binom{|\alpha|}{\alpha} = \binom{\alpha_1 + \alpha_2}{\alpha_1}$. For example, a polynomial of degree 3 is written as

$$\begin{aligned} f(u_1, u_2) = & f_{0,0} + f_{1,0}u_1 + f_{0,1}u_2 + f_{2,0}u_1^2 + 2f_{1,1}u_1u_2 + f_{0,2}u_2^2 \\ & + f_{3,0}u_1^3 + 3f_{2,1}u_1^2u_2 + 3f_{1,2}u_1u_2^2 + f_{0,3}u_2^3. \end{aligned} \quad (25)$$

Example 12. Now consider the case $m = 3$, $d = 3$. Then a polynomial has a representation

$$\begin{aligned} f(u_1, u_2, u_3) = & f_{0,0,0} + f_{1,0,0}u_1 + f_{0,1,0}u_2 + f_{0,0,1}u_3 \\ & + f_{2,0,0}u_1^2 + 2f_{1,1,0}u_1u_2 + 2f_{1,0,1}u_1u_3 + f_{0,2,0}u_2^2 + 2f_{0,1,1}u_2u_3 + f_{0,0,2}u_3^2 \\ & + f_{3,0,0}u_1^3 + 3f_{2,1,0}u_1^2u_2 + 3f_{2,0,1}u_1^2u_3 + 3f_{1,2,0}u_1u_2^2 + 6f_{1,1,1}u_1u_2u_3 \\ & + 3f_{1,0,2}u_1u_3^2 + f_{0,3,0}u_2^3 + 3f_{0,2,1}u_2^2u_3 + 3f_{0,1,2}u_2u_3^2 + f_{0,0,3}u_3^3. \end{aligned}$$

5.2. Compact representation of symmetric tensors

Let \mathcal{T} be a d -th order $m \times \cdots \times m$ symmetric tensor. Recall that the contraction

$$f(\mathbf{u}) = \mathcal{T} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \cdots \bullet_d \mathbf{u} \quad (26)$$

defines a homogenous polynomial of degree d .

The representation (24) of $f(\mathbf{u})$ has the form:

$$f(\mathbf{u}) = \sum_{\alpha \in \mathbb{Z}_+^m, |\alpha|=d} \binom{|\alpha|}{\alpha} f_{\alpha} \mathbf{u}^{\alpha}, \quad (27)$$

because a homogeneous polynomial contains only monomials of degree d . There is a simple relationship between the coefficients of \mathcal{T} and the normalized polynomial coefficients f_{α} .

Example 13. Consider a homogeneous bivariate polynomial (the highest degree homogeneous part from Example 11), given as

$$f(u_1, u_2) = f_{3,0}u_1^3 + 3f_{2,1}u_1^2u_2 + 3f_{1,2}u_1u_2^2 + f_{0,3}u_2^3.$$

Then it is easy to see that (26) becomes

$$f(\mathbf{u}) = \mathcal{T} \bullet_1 \mathbf{u} \bullet_2 \mathbf{u} \bullet_3 \mathbf{u},$$

where the tensor \mathcal{T} is given as

$$\mathcal{T}_{1,1,1} = f_{3,0}, \quad \mathcal{T}_{1,1,2} = f_{2,1}, \quad \mathcal{T}_{1,2,2} = f_{1,2}, \quad \mathcal{T}_{2,2,2} = f_{0,3}.$$

In this case, the matricization of \mathcal{T} with respect to the first mode is equal to

$$\mathcal{T}_{(1)} = \left[\begin{array}{cc|cc} f_{3,0} & f_{2,1} & f_{2,1} & f_{1,2} \\ f_{2,1} & f_{1,2} & f_{1,2} & f_{0,3} \end{array} \right],$$

which exactly corresponds to the rightmost block of the matrix in (11).

Example 13 can be generalized to arbitrary symmetric tensors.

Lemma 4. Let \mathcal{T} be a symmetric tensor and $f(\mathbf{u})$ be the corresponding homogeneous polynomial (given by (26)). Furthermore, for a multi-index

$$\beta \in \{1, \dots, m\}^{\times d},$$

we define the vector $\pi(\beta) \in \mathbb{Z}_+^m$ that counts the number of occurrences of a number in the vector, i.e.

$$\pi(\beta)_i = \#\{\beta_j : \beta_j = i\}. \quad (28)$$

Then the normalized coefficients f_{α} in (24) satisfy the following relation

$$f_{\pi(\beta)} = \mathcal{T}_{\beta},$$

i.e., the coefficients α count the occurrences of indices in \mathcal{T} .

Lemma 4 explains the meaning of the multinomial coefficients in (24): they count the number of occurrences of the same element in a symmetric tensor.

5.3. Quasi-Hankel matrices

Next, we define a structured matrix that will be useful in the proofs. Denote

$$M = \binom{m+d-1}{d-1},$$

which is equal to the number of monomials of a polynomial of degree $d-1$. Suppose that the multidegrees of these monomials are ordered such that

$$\{\alpha^{(1)}, \dots, \alpha^{(M)}\} \subset \mathbb{Z}_+^m. \quad (29)$$

For a polynomial f given as (24), we define the $m \times M$ *quasi-Hankel* matrix² $\mathbf{Q}(f)$ as follows:

$$(\mathbf{Q}(f))_{j,k} = f_{\mathbf{e}_j + \alpha^{(k)}}, \quad (30)$$

where \mathbf{e}_j , $j = 1, \dots, m$, are the unit coordinate vectors.

For example, the matrix $\mathbf{Q}(f)$ for the polynomial in Example 11 is

$$\mathbf{Q}(f) = \left[\begin{array}{c|cc} f_{1,0} & f_{2,0} & f_{1,1} \\ f_{0,1} & f_{1,1} & f_{0,2} \end{array} \middle| \begin{array}{ccc} f_{3,0} & f_{2,1} & f_{1,2} \\ f_{2,1} & f_{1,2} & f_{0,3} \end{array} \right], \quad (31)$$

if the ordering of multidegrees in (29) is chosen as

$$\{(0,0), (1,0), (0,1), (2,0), (1,1), (0,2)\}.$$

Next, the matrix in Example 12 is

$$\mathbf{Q}(f) = \left[\begin{array}{c|ccc|ccc|ccc} f_{1,0,0} & f_{2,0,0} & f_{1,1,0} & f_{1,0,1} & f_{3,0,0} & f_{2,1,0} & f_{2,0,1} & f_{1,2,0} & f_{1,1,1} & f_{1,0,2} \\ f_{0,1,0} & f_{1,1,0} & f_{0,2,0} & f_{0,1,1} & f_{2,1,0} & f_{1,2,0} & f_{1,1,1} & f_{0,3,0} & f_{0,2,1} & f_{0,1,2} \\ f_{0,0,1} & f_{1,0,1} & f_{0,1,1} & f_{0,0,2} & f_{2,0,1} & f_{1,1,1} & f_{1,0,2} & f_{0,2,1} & f_{0,1,2} & f_{0,0,3} \end{array} \right], \quad (32)$$

where the ordering of multidegrees is chosen as:

$$\{(0,0,0), (1,0,0), (0,1,0), (0,0,1), (2,0,0), (1,1,0), (1,0,1), (0,2,0), (0,1,1), (0,0,2)\}.$$

We also mention that the matrix $\mathbf{Q}(f)$ is exactly the structured matrix considered in [28, Proposition 4.1].

6. Proofs of main results

In this section, we provide the proofs of the results from the previous sections.

²The term “quasi-Hankel” was introduced in [35]

6.1. Compressed representation of \mathcal{Q}

Proof of Proposition 2. First we define the selection matrix $\mathbf{S} \in \mathbb{R}^{\delta \times M}$ as follows. Assume that the multidegrees of monomials are ordered as in (29). Next, we define the set of multi-indices

$$\{\beta^{(1)}, \dots, \beta^{(\delta)}\}$$

as follows: $\beta^{(1)} = 0$, and for each $j \in \{1, \dots, d-1\}$ we define $\delta_{j-1} = \sum_{s=0}^{j-1} m^s$, and

$$\{\beta^{(\delta_{j-1}+1)}, \dots, \beta^{(\delta_{j-1}+m^j)}\} \subset \{1, \dots, m\}^{\times j}$$

contain the ordered elements of $\{1, \dots, m\}^{\times j}$ in the order of vectorization. For example, for $m = 2$ and $d = 3$ we have the ordered set

$$\{\beta^{(1)}, \dots, \beta^{(\delta)}\} = \{0, 1, 2, (1, 1), (2, 1), (1, 2), (2, 2)\}.$$

For $m = 3$ and $d = 3$ we get

$$\{\beta^{(1)}, \dots, \beta^{(\delta)}\} = \{0, 1, 2, 3, (1, 1), (2, 1), (3, 1), (1, 2), (2, 2), (3, 2), (1, 3), (2, 3), (3, 3)\}.$$

It is easy to see that the coefficients $\beta^{(i)}$ correspond to the indices of the elements in the matrix of unfoldings $\Psi(f)$.

Next for a multi-index

$$\beta \in \{1, \dots, m\}^{\times j},$$

we define $\pi(\beta) \in \mathbb{Z}_+^m$ as in (28). Finally, we define the matrix $\mathbf{S} \in \mathbb{R}^{\delta \times M}$ as follows

$$\mathbf{S}_{i,j} = \begin{cases} 1, & \text{if } i = j = 1, \\ 1, & \text{if } \pi(\beta^{(i)}) = \alpha^{(j)}, \\ 0, & \text{otherwise.} \end{cases}$$

By construction, we have that \mathbf{S} contains only one nonzero element in each row. Moreover, the matrix has full column rank and all columns are orthogonal to each other.

From Lemma 4 it is immediate to see that

$$\Psi(f) = \mathbf{Q}(f)\mathbf{S}^\top.$$

Next, define the tensor $\mathcal{C} \in \mathbb{R}^{n \times m \times M}$ by stacking the slices

$$\mathcal{Q}_{i,:,:} := \mathbf{Q}(f_i),$$

and we have that $\mathcal{Q} = \mathcal{C} \bullet_3 \mathbf{S}$ by construction.

Finally, let us construct the matrix $\mathbf{B} \in \mathbb{R}^{M \times r}$ as follows:

$$\mathbf{B}_{j,k} = |\alpha^{(j)}|(\mathbf{u}^{(k)})^{\alpha^{(j)}},$$

which is a multivariate Vandermonde matrix. Then we have that

$$\mathbf{A}_{i,k} = |\pi(\beta^{(i)})|(\mathbf{u}^{(k)})_{(\beta^{(i)})_1}(\mathbf{u}^{(k)})_{(\beta^{(i)})_2} \cdots (\mathbf{u}^{(k)})_{(\beta^{(i)})_m} = |\pi(\beta^{(i)})|(\mathbf{u}^{(k)})^{\pi(\beta^{(i)})},$$

hence $\mathbf{A} = \mathbf{S}\mathbf{B}$. □

Proof of Proposition 3.

1. First, since $\mathcal{Q} = \mathcal{C} \bullet_3 \mathbf{S}$ and \mathbf{S} has full column rank, the CP ranks of \mathcal{Q} and \mathcal{C} coincide, and there is a one-to-one correspondence between their decompositions. Second, we have

$$\mathcal{J} = \mathcal{Q} \bullet_3 \mathbf{A}^\top = (\mathcal{C} \bullet_3 \mathbf{S}) \bullet_3 (\mathbf{B}^\top \mathbf{S}^\top) = \mathcal{C} \bullet_3 (\mathbf{B}^\top \mathbf{S}^\top \mathbf{S}).$$

By orthogonality of the columns of \mathbf{S} , the matrix $\mathbf{S}^\top \mathbf{S}$ is a nonsingular diagonal matrix. Moreover, \mathbf{B}^\top has full column rank, hence the ranks of \mathcal{J} and \mathcal{C} coincide, and their decompositions are equivalent.

2. For a generic choice of points, if $N \geq M$, the multivariate Vandermonde matrix has rank M , since all the multidegrees $\alpha^{(j)}$ are distinct.

□

6.2. Structure and CP rank

Before proving Proposition 4, we show that the tensor \mathcal{C} , in fact, contains all the coefficients of the Jacobian of \mathbf{f} .

Lemma 5. *Consider a scaled version of the matrix $\mathbf{Q}(f)$ given by*

$$\mathbf{C}(f) = \mathbf{Q}(f) \begin{bmatrix} 1 & & & \\ & 2I_{M_1} & & \\ & & \ddots & \\ & & & dI_{M_{d-1}} \end{bmatrix},$$

where I_K is the $K \times K$ identity, and $M_k = \binom{m+k-1}{k}$. Then it holds that the j -th row of $\mathbf{C}(f)$ contains the coefficients of $\frac{\partial f}{\partial u_j}$.

Proof. It is easy to see the following

$$\begin{aligned} \frac{\partial f}{\partial u_j}(u_1, \dots, u_m) &= \sum_{\alpha_1, \dots, \alpha_m=0}^{|\alpha| \leq d, \alpha_j \neq 0} \alpha_1 \frac{(\alpha_1 + \dots + \alpha_m)!}{\alpha_1! \dots \alpha_m!} f_{\alpha_1, \dots, \alpha_m} u_1^{\alpha_1} \dots u_{j-1}^{\alpha_{j-1}} u_j^{\alpha_j-1} u_{j+1}^{\alpha_{j+1}} \dots u_m^{\alpha_m} \\ &= \sum_{\alpha_1, \dots, \alpha_m=0}^{|\alpha| \leq d-1} (|\alpha| + 1) \binom{|\alpha|}{\alpha} f_{\alpha_1, \dots, \alpha_{j-1}, \alpha_j+1, \alpha_{j+1}, \dots, \alpha} \mathbf{u}^\alpha, \end{aligned} \tag{33}$$

which completes the proof. □

Example 14. *Consider an example $m = 2$ and $d = 3$. Take a polynomial in (25). Let the partial derivatives be*

$$\begin{aligned} \frac{\partial f}{\partial u_1} &= f_{1,0} + 2f_{2,0}u_1 + 2f_{1,1}u_2 + 3f_{3,0}u_1^2 + 6f_{2,1}u_1u_2 + 3f_{1,2}u_2^2, \\ \frac{\partial f}{\partial u_2} &= f_{0,1} + 2f_{1,1}u_1 + 2f_{0,2}u_2 + 3f_{2,1}u_1^2 + 6f_{1,2}u_1u_2 + 3f_{0,3}u_2^2. \end{aligned}$$

If we choose (for the derivatives) an ordering of multidegrees in (29) as

$$\{(0,0), (1,0), (0,1), (2,0), (1,1), (0,2)\},$$

then the slice $\mathbf{C}(f)$ has the following form

$$\mathbf{C}(f) = \begin{bmatrix} f_{1,0} & 2f_{2,0} & 2f_{1,1} & 3f_{3,0} & 3f_{2,1} & 3f_{1,2} \\ f_{0,1} & 2f_{1,1} & 2f_{0,2} & 3f_{2,1} & 3f_{1,2} & 3f_{0,3} \end{bmatrix}.$$

We see that the rows of $\mathbf{C}(f)$ are exactly the coefficients of the partial derivatives in the multinomial basis.

Proof of Proposition 4. The \Rightarrow follows from Lemma 1.

Let us prove the \Leftarrow part. Assume that \mathcal{Q} has the decomposition

$$\mathcal{C} = \mathbf{w} \circ \mathbf{v} \circ \mathbf{y}.$$

First, since by Lemma 5, the tensor \mathcal{C} contains all the coefficients of the derivatives, we have that there exists a polynomial $\tilde{f}(\mathbf{u})$ such that

$$\nabla f_k(\mathbf{u}) = (\mathbf{w})_k \nabla \tilde{f}(\mathbf{u}).$$

Since the polynomials $f_k(\mathbf{u})$ do not have constant terms, we have that

$$\mathbf{f}(\mathbf{u}) = \mathbf{w} \tilde{f}(\mathbf{u}),$$

where

$$\mathbf{Q}(\tilde{f}) = \mathbf{v} \mathbf{y}^\top.$$

Hence, by [28, Proposition 4.1], the polynomial \tilde{f} has necessarily the form $\tilde{f}(\mathbf{u}) = g(\mathbf{v}^\top \mathbf{u})$, which completes the proof. \square

7. Conclusions and Perspectives

We discovered and studied a useful link between two tensorization approaches for decoupling multivariate polynomials [19, 7]: the tensor of Jacobian matrices [7] can be obtained by multiplying the coefficient-based tensor [19] by a Vandermonde-like matrix. As revealed by this connection, the two approaches have similar fundamental properties, such as equal tensor rank.

In addition, although tensors often have unique CP decompositions, we extended the uniqueness property for the decoupling problem even further, as it is an essential assumption for the correct working of the algorithms. By imposing the underlying structure on the factors of the CP decompositions, we reduced the solution space and thus provided identifiability of an even larger class of problems.

The obtained insights shed a light on the applicability of the two approaches. First, properties can now be transferred from one approach to the other. For example, uniqueness aspects in one of the settings would lead to uniqueness aspects in the other. Second, as the underlying principles of the approaches are different, each of them is suitable for solving different variations and generalizations of the main decoupling problem. Thus, understanding the advantages and disadvantages of the approaches helps choosing the

more suitable one. We mention that, for (differentiable) non-polynomial functions, the approach based on Jacobian matrices would be more appropriate, as it only uses evaluations of the derivatives of the functions. On the other hand, when the region of interest is unclear and it is difficult to generate representative sampling points, the coefficient-based approach should be chosen. Finally, in the single polynomial case, with the structured data fusion approach, non-uniqueness of the matrix decomposition of the degenerate tensors can be avoided.

Future work on the decoupling problem includes imposing the polynomial structure in the Jacobian tensor, by parametrizing the \mathbf{H} factor of the CP decomposition by the coefficients (of the partial derivatives) of the polynomials. Furthermore, we will study the structured CP decompositions of noisy tensors, based on recent results of the unstructured case [36]. Another interesting problem is computing decompositions of tensors based on incomplete function descriptions, e.g., when some of the coefficients are missing or unreliable. For the latter, the coefficient-based approach seems more appropriate.

8. Acknowledgements

This work was supported in part by ERC AdG-2013-320594 grant “DECODA”, by the Fund for Scientific Research (FWO-Vlaanderen), by the Flemish Government (Methusalem), the Belgian Government through the Inter-university Poles of Attraction (IAP VII) Program, by the ERC Advanced Grant SNLSID under contract 320378, and by FWO projects G.0280.15N and G.0901.17N.

References

References

- [1] J. Alexander, A. Hirschowitz, Polynomial interpolation in several variables, *J. Algebraic Geom.* 4 (2) (1995) 201–222.
- [2] A. Iarrobino, V. Kanev, *Power Sums, Gorenstein Algebras, and Determinantal Loci*, Vol. 1721 of *Lecture Notes in Mathematics*, Springer, 1999.
- [3] J. M. Landsberg, *Tensors: Geometry and Applications*, Vol. 128 of *Grad. Stud. Math.*, American Mathematical Society, Providence, RI, 2012.
- [4] A. Białynicki-Birula, A. Schinzel, Representations of multivariate polynomials as sums of polynomials in linear forms, *Colloquium Mathematicum* 112 (2) (2008) 201–233.
- [5] A. Schinzel, On a decomposition of polynomials in several variables, *J. de Théorie des Nombres de Bordeaux* 14 (2) (2002) 647–666.
- [6] E. Carlini, J. Chipalkatti, On Waring’s problem for several algebraic forms, *Comment. Math. Helv.* 78 (2003) 494–517.
- [7] P. Dreesen, M. Ishteva, J. Schoukens, Decoupling multivariate polynomials using first-order information, *SIAM. J. Matrix Anal. Appl.* 36 (2) (2015) 864–879.
- [8] K. Tiels, J. Schoukens, From coupled to decoupled polynomial representations in parallel Wiener-Hammerstein models, in: *Proc. 52nd IEEE Conf. Decis. Control (CDC)*, Florence, Italy, 2013, pp. 4937–4942.
- [9] B. F. Logan, L. A. Shepp, [Optimal reconstruction of a function from its projections](https://doi.org/10.1215/S0012-7094-75-04256-8), *Duke Math. J.* 42 (4) (1975) 645–659. doi:10.1215/S0012-7094-75-04256-8. URL <http://dx.doi.org/10.1215/S0012-7094-75-04256-8>
- [10] V. Lin, A. Pinkus, Fundamentality of ridge functions, *Journal of Approximation Theory* 75 (3) (1993) 295 – 311. doi:<http://dx.doi.org/10.1006/jath.1993.1104>.
- [11] K. I. Oskolkov, On representations of algebraic polynomials as a sum of plane waves, *Serdica Mathematical Journal* (2002) 379–390.

- [12] Y. Shin, J. Ghosh, Ridge polynomial networks, *IEEE Transactions on Neural Networks* 6 (3) (1995) 610–622. doi:10.1109/72.377967.
- [13] P. Dreesen, M. Schoukens, K. Tiels, J. Schoukens, Decoupling static nonlinearities in a parallel Wiener-Hammerstein system: A first-order approach, in: *Proc. 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2015)*, Pisa, Italy, 2015, pp. 987–992.
- [14] P. Dreesen, A. Fakhrizadeh Esfahani, J. Stoev, K. Tiels, J. Schoukens, Decoupling nonlinear state-space models: case studies, in: P. Sas, D. Moens, A. van de Walle (Eds.), *International Conference on Noise and Vibration (ISMA2016) and International Conference on Uncertainty in Structural Dynamics (USD2016)*, Leuven, Belgium, 2016, pp. 2639–2646.
- [15] M. Schoukens, Y. Rolain, Cross-term elimination in parallel Wiener systems using a linear input transformation, *IEEE Transactions on Instrumentation and Measurement* 61 (3) (2012) 845–847. doi:10.1109/TIM.2011.2174851.
- [16] M. Schoukens, K. Tiels, M. Ishteva, J. Schoukens, Identification of parallel Wiener-Hammerstein systems with a decoupled static nonlinearity, in: *Proceedings of 19th IFAC World Congress*, Cape Town (South Africa), August 24–29, 2014, 2014.
- [17] A. Fakhrizadeh Esfahani, P. Dreesen, K. Tiels, J.-P. Noël, J. Schoukens, G. Kerschen, Polynomial state-space model decoupling for the identification of hysteretic systems, in: *20th IFAC World Congress 2017*, Toulouse, France (Accepted), 2017.
- [18] K. Tiels, J. Schoukens, From coupled to decoupled polynomial representations in parallel Wiener-Hammerstein models, in: *52nd IEEE Conference on Decision and Control*, Florence, Italy, December 10–13, 2013, 2013, pp. 4937–4942.
- [19] A. Van Mulders, L. Vanbeylen, K. Usevich, Identification of a block-structured model with several sources of nonlinearity, in: *Proceedings of the 14th European Control Conference (ECC 2014)*, 2014, pp. 1717–1722. doi:10.1109/ECC.2014.6862455.
- [20] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, L. De Lathauwer, Tensorlab 3.0, available online, Mar. 2016. URL: <http://www.tensorlab.net/> (2016).
- [21] C. A. Andersson, R. Bro, The N-way toolbox for MATLAB, *Chemometrics & Intelligent Laboratory Systems* 52 (2000) 1–4, <http://www.models.life.ku.dk/source/nwaytoolbox/>.
- [22] B. W. Bader, T. G. Kolda, et al., MATLAB tensor toolbox version 2.5, available online, January 2012. URL: <http://www.sandia.gov/~tgkolda/TensorToolbox/> (2012).
- [23] O. Debals, L. De Lathauwer, Stochastic and deterministic tensorization for blind signal separation, in: *Proc. 12th International Conference on Latent Variable Analysis and Signal Separation (LVA-ICA 2015)*, Vol. 9237 of Lecture Notes in Computer Science, Springer, 2015, pp. 3–13.
- [24] J. Carroll, J. Chang, Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition, *Psychometrika* 35 (3) (1970) 283–319.
- [25] R. A. Harshman, Foundations of the PARAFAC procedure: Model and conditions for an “explanatory” multi-mode factor analysis, *UCLA Working Papers in Phonetics* 16 (1) (1970) 1–84.
- [26] T. G. Kolda, B. W. Bader, Tensor decompositions and applications, *SIAM Rev.* 51 (3) (2009) 455–500.
- [27] P. Comon, Y. Qi, K. Usevich, A polynomial formulation for joint decomposition of symmetric tensors of different orders, in: E. Vincent, A. Yeredor, Z. Koldovský, P. Tichavský (Eds.), *Latent Variable Analysis and Signal Separation*, Vol. 9237 of Lecture Notes in Computer Science, Springer International Publishing, 2015, pp. 22–30. doi:10.1007/978-3-319-22482-4_3. URL http://dx.doi.org/10.1007/978-3-319-22482-4_3
- [28] P. Comon, Y. Qi, K. Usevich, Identifiability of an X-rank decomposition of polynomial maps, *Tech. rep.*, submitted to *SIAM Journal on Applied Algebra and Geometry*. Available from <http://arxiv.org/abs/1603.01566>. (2016).
- [29] J. Brachat, P. Comon, B. Mourrain, E. Tsigaridas, Symmetric tensor decomposition, *Lin. Alg. Appl.* 433 (11) (2010) 1851–1872.
- [30] K. Batselier, N. Wong, Symmetric tensor decomposition by an iterative eigendecomposition algorithm, *J. Comput. Appl. Math.* 308 (2016) 69–82.
- [31] P. Comon, G. H. Golub, L.-H. Lim, B. Mourrain, Symmetric tensors and symmetric tensor rank, *SIAM. J. Matrix Anal. Appl.* 30 (3) (2008) 1254–1279.
- [32] K. Ranestad, F.-O. Schreyer, Varieties of sums of powers, *J. Reine Angew. Math.* 525 (2000) 147–181.
- [33] L. Sorber, M. Van Barel, L. De Lathauwer, Structured data fusion, *IEEE J. Sel. Top. Signal Process.* 9 (4) (2015) 586–600.
- [34] A. Van Mulders, J. Schoukens, L. Vanbeylen, Identification of systems with localised nonlinearity: From state-space to block-structured models, *Automatica* 49 (5) (2013) 1392 – 1396.

doi:<http://dx.doi.org/10.1016/j.automat.2013.01.052>.

URL <http://www.sciencedirect.com/science/article/pii/S0005109813000538>

- [35] B. Mourrain, V. Y. Pan, Multivariate polynomials, duality, and structured matrices, *Journal of complexity* 16 (1) (2000) 110–180.
- [36] G. Hollander, P. Dreesen, M. Ishteva, J. Schoukens, Weighted tensor decomposition for approximate decoupling of multivariate polynomials (2016). [arXiv:arXiv:1601.07800](#).